

XP008008413

sub-

Q-32. Detection of Dechlorinating Bacteria in Groundwater and Soils from Waste Sites Contaminated with PCE and TCE

M.A. ELBERSON¹, R.C. EBERSOLE¹, R.M. YOUNG², J.A. TABINOWSKI¹, M.P. PERRY², S. FAHNESTOCK², E.R. HENDRICKSON¹,
¹Microbial and Industrial Microbiology, Life Science Enterprise, Du Pont Co., Glasgow Site, Newark, DE; ²Du Pont Co., Experimental Station, Wilmington, DE

Chlorinated ethenes, perchloroethene (PCE) and trichloroethene (TCE), are contaminants of groundwater at many hazardous waste sites. These solvents are both toxic and suspected carcinogens. Recently, *Dehalococcoides ethenogenes*, an organism described Maymo-Gatell et al. (Science 276, 1568-1571 (1997)), was shown to dechlorinate PCE and TCE by removing all the chlorine atoms, through a process known as dehalorespiration, to form ethene. In addition, other microorganisms, such as, *Dehalobacter restrictus* and *Dehalospirillum multivorans*, have demonstrated partial dechlorination of chlorinated ethenes. Based on their 16S rRNA gene sequences, we have developed a specific PCR assays to detect each of these organisms. Using these assays, the presence of these dehalorespiring organisms has been demonstrated in microcosms inoculated with groundwater or soil from four waste sites contaminated with PCE and TCE. The PCR amplicons were sequenced to ver-

US 095489980YP1



Creation date: 10-16-2003
Indexing Officer: MBEYENE - MELKAM BEYENE
Team: OIPEScanning
Dossier: 09548998

Legal Date: 02-09-2002

No.	Doccode	Number of pages
1	SRNT	34

Total number of pages: 34

Remarks:

Order of re-scan issued on



A DOCPHOENIX

APPL PARTS

IMIS
Internal Misc. Paper

LET.
Misc. Incoming Letter

371P
PCT Papers in a 371 Application

A...
Amendment Including Elections

ABST
Abstract

ADS
Application Data Sheet

AF/D
Affidavit or Exhibit Received

APPENDIX
Appendix

ARTIFACT
Artifact

BIB
Bib Data Sheet

CLM
Claim

COMPUTER
Computer Program Listing

CRFL
All CRF Papers for Backfile

DIST
Terminal Disclaimer Filed

DRW
Drawings

FOR
Foreign Reference

FRPR
Foreign Priority Papers

IDS
IDS Including 1449

NPL
Non-Patent Literature

OATH
Oath or Declaration

PET.
Petition

RETMAIL
Mail Returned by USPS

SEQLIST
Sequence Listing

SPEC
Specification

SPEC NO
Specification Not in English

TRNA
Transmittal New Application

OUTGOING

CTMS
Misc. Office Action

1449
Signed 1449

892
892

ABN
Abandonment

APDEC
Board of Appeals Decision

APEA
Examiner Answer

CTAV
Count Advisory Action

CTEQ
Count Ex parte Quayle

CTFR
Count Final Rejection

Duplex
02/09/02

CTNF
Count Non-Final

CTRS
Count Restriction

EXIN
Examiner Interview

M903
DO/EO Acceptance

M905
DO/EO Missing Requirement

NFDR
Formal Drawing Required

NOA
Notice of Allowance

PETDEC
Petition Decision

INCOMING

AP.B
Appeal Brief

C.AD
Change of Address

N/AP
Notice of Appeal

PA..
Change in Power of Attorney

REM
Applicant Remarks in Amendment

XT/
Extension of Time filed separate

Internal

SRNT
Examiner Search Notes

CLMPTO
PTO Prepared Complete Claim Set

ECBOX
Evidence Copy Box Identification

WCLM
Claim Worksheet

WFEE
Fee Worksheet

File Wrapper

FWCLM
File Wrapper Claim

IIFW
File Wrapper Issue Information

SRFW
File Wrapper Search Info

59984

Hart, Edward

From: Loeb, Bronwen
Sent: Friday, February 08, 2002 6:32 PM
To: Hart, Edward
Subject: sequence search request 09/548,998

Please search SEQ ID No. 8 from 09/548,998 in all the nucleotide databases except the pending databases.

Thank you.

Bronwen Loeb, PhD
AU 1636
703-605-1197
CM1 12D12
Mailbox 11E-12

Edward Hart
Technical Support
Seq. ID No. 8
09/548,998

2/9/02
AB5504
1-NA

<u>Seq ID No</u>	<u># nt</u>
1	24
2	1377
3	1378
4	1377
5	1377
6	1377
8	47
30	51
34	49

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 1344.17 Seconds

(without alignments)
576.837 Million cell updates/sec

Title: US-09-548-998A-8

Perfect score: 47
Sequence: 1 tgtgtggygcycacatawgtl.....ytaaagccgyaaggygcttg 47

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inu:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inu:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	44.2	94.0	47	6	AX039541	AX039541 Sequence
2	44.2	94.0	1212	6	AX039535	AX039535 Sequence
3	44.2	94.0	1212	6	AX039537	AX039537 Sequence
4	44.2	94.0	1212	6	AX039538	AX039538 Sequence
5	44.2	94.0	1212	6	AX039539	AX039539 Sequence
6	44.2	94.0	1322	1	AF357918	AF357918 Dehalococ
7	44.2	94.0	1322	1	UBA249262	AJ249262 bacterium
8	44.2	94.0	1335	6	AX039536	AX039536 Sequence
9	44.2	94.0	1421	1	AF230641	AF230641 Bacterium
10	44.2	94.0	1434	1	AF004928	AF004928 Dehalococ
11	44.2	94.0	1443	6	AX039540	AX039540 Sequence
12	26.4	56.2	87286	2	AC014436	AC014436 Drosophil
13	26.4	56.2	170914	3	AC010843	AC010843 Drosophil
14	26.4	56.2	185810	3	AC010995	AC010995 Drosophil
15	26.4	56.2	298839	3	AE003492	AE003492 Drosophil
16	24.6	52.3	318503	2	AC074166	AC074166 Mus muscu
17	24.4	51.9	3215	5	ORNFGLUD	L34080 Oreochromis
18	24.4	51.9	3215	5	ORNFGLUD	L46366 Oreochromis
19	24.4	51.9	3391	5	ORNFGLUD	L34036 Oreochromis
20	24.4	51.9	59843	8	AP000816	AF389400 Danio rer
21	24.2	51.5	2421	5	AF389400	U41288 Allomyces m
22	24.2	51.5	57473	8	AMU41288	AF005746 Unculture
23	23.8	50.6	1355	1	AF005746	AF005746 Unculture
24	23.6	50.2	410	1	AF010041	AF010041 Unidentif
25	23.4	49.8	110000	2	LMFLCHR32_14	Continuation (15 o
26	23.3	48.9	184841	2	AC079226	AC079226 Homo sapi
27	22.8	48.5	595	3	AF155610	AF155610 Polistes
28	22.8	48.5	119184	1	D90909	D90909 Synecocyst
29	22.8	48.5	170139	2	AC073727	AC073727 Mus muscu
30	22.8	48.5	195470	2	AC079367	AC079367 Mus muscu
31	22.8	48.5	235411	2	AC084066	AC084066 Mus muscu
32	22.6	48.1	173271	2	AC023919	AC023919 Homo sapi
33	22.4	47.7	160628	2	AC091243	AC091243 Trypanoso
34	22.4	47.7	184457	1	AF079317	AF079317 Sphingomo
35	22.2	47.2	2557	8	SCYDL063C	274111 S. cerevisia
36	22.2	47.2	142047	9	AC026708	AC026708 Homo sapi
37	22.2	47.2	160025	2	AC018746	AC018746 Homo sapi
38	22.2	47.2	162712	2	AC015677	AC015677 Homo sapi
39	22.2	47.2	165428	2	AP001099	AP001099 Homo sapi
40	22.2	47.2	172720	2	AC091886	AC091886 Homo sapi
41	22.2	47.2	179391	2	AP001283	AP001283 Homo sapi
42	22.2	47.2	203450	2	AP002009	AP002009 Homo sapi
43	22.2	47.2	208921	2	AP001911	AP001911 Homo sapi
44	22	46.8	1420	1	AF027031	AF027031 Unidentif
45	22	46.8	1420	1	AF027032	AF027032 Unidentif

ALIGNMENTS

RESULT 1
AX039541

LOCUS AX039541 47 bp DNA
DEFINITION Sequence 8 from Patent WO0063443.
ACCESSION AX039541
VERSION AX039541.1 GI:11229563

PAT 18-NOV-2000

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE

Journal
Patent: WO 0063443-A 8 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
Source
1..47
/organism="synthetic construct"
/db_xref="taxon:32630"

unsure /note="CONSENSUS"
5 /note="R-A/G"
11 /note="Y-C/T"
18 /note="W-A/T"
21 /note="Y-C/T"
28 /note="Y-T/C"
37 /note="Y-T/C"
42 /note="Y-T/C"
unsure /note="Y-C/T"
9 a 6 c 15 g 10 t 7 others
BASE COUNT
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 47;
Best Local Similarity 100.0%; Pred. No. 4.2e-08;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgggcygacatawgttgctcaytaagccgyaaggygcttg 47
|||||
Db 1 TGTGTGGGCGYACATATGTTGTTCACTAAAGCCGTAAGGCGCTTG 47

RESULT 2
AX039535 1212 bp DNA PAT 18-NOV-2000
LOCUS
DEFINITION Sequence 2 from Patent WO0063443.
ACCESSION AX039535
VERSION AX039535.1 GI:11229557
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.

REFERENCE
AUTHORS 1 (bases 1 to 1212)
Hendrickson, E.R. and Ebersole, R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria

JOURNAL
E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
source 1.1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"

BASE COUNT 321 a 252 c 369 g 269 t 1 others
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgggcygacatawgttgctcaytaagccgyaaggygcttg 47
|||||
Db 140 TGTGTGGGCGYACATATGTTGTTCACTAAAGCCGTAAGGCGCTTG 186

RESULT 3
AX039537 1212 bp DNA PAT 18-NOV-2000
LOCUS
DEFINITION Sequence 4 from Patent WO0063443.
ACCESSION AX039537
VERSION AX039537.1 GI:11229559
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.

REFERENCE 1 (bases 1 to 1212)
AUTHORS Hendrickson, E.R. and Ebersole, R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria
JOURNAL Patent: WO 0063443-A 4 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source 1.1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"

BASE COUNT 318 a 253 c 372 g 269 t
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgggcygacatawgttgctcaytaagccgyaaggygcttg 47
|||||
Db 140 TGTGTGGGCGYACATATGTTGTTCACTAAAGCCGTAAGGCGCTTG 186

RESULT 4
AX039538 1212 bp DNA PAT 18-NOV-2000
LOCUS
DEFINITION Sequence 5 from Patent WO0063443.
ACCESSION AX039538
VERSION AX039538.1 GI:11229560
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.

REFERENCE
AUTHORS 1 (bases 1 to 1212)
Hendrickson, E.R. and Ebersole, R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria

JOURNAL
E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
source 1.1212
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"

BASE COUNT 320 a 253 c 371 g 268 t
ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

OY 1 tgtgtgggcygacatawgttgctcaytaagccgyaaggygcttg 47
|||||
Db 140 TGTGTGGGCGYACATATGTTGTTCACTAAAGCCGTAAGGCGCTTG 186

RESULT 5
AX039539 1212 bp DNA PAT 18-NOV-2000
LOCUS
DEFINITION Sequence 6 from Patent WO0063443.
ACCESSION AX039539
VERSION AX039539.1 GI:11229561
KEYWORDS
SOURCE
ORGANISM
Dehalococcoides ethenogenes.
Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.

REFERENCE
AUTHORS 1 (bases 1 to 1212)
Hendrickson, E.R. and Ebersole, R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating bacteria

JOURNAL
Patent: WO 0063443-A 6 26-OCT-2000;

E.I. DU PONT DE NEMOURS AND COMPANY (US)

FEATURES
source
1. 1312
Location/Qualifiers

/organism="Dehalococcoides ethenogenes"

BASE COUNT 320 a 255 c 371 g 266 t

ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1212;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;QY 1 tgtgtggygcacatawgttggttcaytaaacgcgyaagygcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 TGTGTGGGGCCGACATAGTGTGCTCACTAAAGCCGTAAGTGTG 186

RESULT 6

AF357918

LOCUS 1322 bp DNA BCT 03-MAY-2001

DEFINITION Dehalococcoides sp. FL2 16S ribosomal RNA gene, partial sequence.

ACCESSION AF357918

VERSION AF357918.1 GI:13936862

KEYWORDS

ORGANISM

Dehalococcoides sp. FL2.

Dehalococcoides sp. FL2.

Bacteria; Green non-sulfur bacteria; Dehalococcoides group;

Dehalococcoides.

REFERENCE

1 (bases 1 to 1322)

Loeffler, F.E.

Dehalococcoides sp. strain FL2 is the trichloroethene-dechlorinating population in a defined co-culture consisting of three distinct populations

Unpublished

JOURNAL

2 (bases 1 to 1322)

REFERENCE

Loeffler, F.E.

AUTHORS

Direct Submission

TITLE

Submitted (08-MAR-2001) Environmental Engineering, Georgia

JOURNAL

Institute of Technology, 200 Bobby Dodd Way, Atlanta, GA 30332, USA

FEATURES

Location/Qualifiers

1. 1322

/organism="Dehalococcoides sp. FL2"

/strain="FL2"

/db_xref="taxon:158193"

<1. >1322

/product="16S ribosomal RNA"

BASE COUNT 347 a 271 c 401 g 302 t 1 others

ORIGIN

Query Match 94.0%; Score 44.2; DB 1; Length 1322;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;QY 1 tgtgtggygcacatawgttggttcaytaaacgcgyaagygcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 159 TGTGTGGGGCCGACATAGTGTGCTCACTAAAGCCGTAAGCGCTTG 205

RESULT 7

UBA249262

LOCUS 1322 bp DNA BCT 28-AUG-2000

DEFINITION bacterium DCEH2 16S rRNA gene.

ACCESSION AJ249262

VERSION AJ249262.1 GI:9944249

KEYWORDS 16S ribosomal RNA; 16S rRNA gene.

SOURCE

bacterium DCEH2.

ORGANISM

Bacteria.

REFERENCE 1 (bases 1 to 1322)

Windfuhr, C., Mau, M., Scholz-Muramatsu, H. and Diekert, G.

Anaerobic reductive dechlorination of chlorinated ethenes with an

enriched mixed culture - physiological characterization and community analysis

Unpublished

2 (bases 1 to 1322)

Windfuhr, C.

Direct Submission

Submitted (31-AUG-1999) Windfuhr C., Institute for Sanitary

Engineering, Department of Biology, University of Stuttgart,

Bandtaele 2, 70569 Stuttgart, 70569, GERMANY

Location/Qualifiers

1. 1322

/organism="bacterium DCEH2"

/db_xref="taxon:103252"

/country="Germany"

/clone="DCEH2"

/note="clone from dechlorinating enrichment mixed culture"

1. 1322

/gene="16S rRNA"

/product="16S ribosomal RNA"

1. 1322

/gene="16S rRNA"

BASE COUNT 346 a 280 c 399 g 294 t 3 others

ORIGIN

Query Match 94.0%; Score 44.2; DB 1; Length 1322;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;QY 1 tgtgtggygcacatawgttggttcaytaaacgcgyaagygcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 154 TGTGTGGGGCCGACATAGTGTGCTCACTAAAGCCGTAAGCGCTTG 200

RESULT 8

AX039536

LOCUS 1335 bp DNA PAT 18-NOV-2000

DEFINITION Sequence 3 from Patent WO0063443.

ACCESSION AX039536

VERSION AX039536.1 GI:11229558

KEYWORDS

ORGANISM

Dehalococcoides ethenogenes.

Dehalococcoides ethenogenes.

Bacteria; Green non-sulfur bacteria; Dehalococcoides group;

Dehalococcoides.

1 (bases 1 to 1335)

Hendrickson, E.R. and Ebersole, R.C.

Nucleic acid fragments for the identification of dechlorinating bacteria

Patent: WO 0063443-A 3 26-OCT-2000;

E.I. DU PONT DE NEMOURS AND COMPANY (US)

Location/Qualifiers

1. 1335

/organism="Dehalococcoides ethenogenes"

/db_xref="taxon:61435"

BASE COUNT 348 a 288 c 403 g 296 t

ORIGIN

Query Match 94.0%; Score 44.2; DB 6; Length 1335;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;QY 1 tgtgtggygcacatawgttggttcaytaaacgcgyaagygcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 TGTGTGGGGCCGACATAGTGTGCTCACTAAAGCCGTAAGTGTG 186

RESULT 9

AF230641

LOCUS 1421 bp DNA BCT 05-DEC-2000

DEFINITION Bacterium CBDB1 16S ribosomal RNA gene, partial sequence.

ACCESSION AF230641

VERSION AF230641.1 GI:11545272
KEYWORDS
SOURCE bacterium CBDB1.
ORGANISM bacterium CBDB1.
REFERENCE 1 (bases 1 to 1421)
AUTHORS Adrian, L., Szezyk, U., Wecke, J. and Gortsch, H.
TITLE Bacterial dehalorespiration with chlorinated benzenes
JOURNAL Nature 408 (6812), 580-583 (2000)
MEDLINE 21003821
REFERENCE 2 (bases 1 to 1421)
AUTHORS Adrian, L., Tran, T., Szezyk, U. and Goerisch, H.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2000) FG Technische Biochemie, TU Berlin, Seestr.
13, Berlin 13353, Germany
FEATURES
source location/Qualifiers
1. 1421
/organism="bacterium CBDB1"
/strain="CBDB1"
/db_xref="taxon:114835"
/note="tentatively related to Dehalococcoides ethenogenes"
<1. .>1421
/product="16S ribosomal RNA"
BASE COUNT 368 a 307 c 435 g 311 t
ORIGIN
Query Match 94.0%; Score 44.2; DB 1; Length 1421;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggtrtgccgacatagtggttcaytaagccgaaagggcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 TGTGTTGGGCGGACATATGTTGTTCTACTTAAGCCGTAAGCGCTTG 186
RESULT 10
AF004928 1434 bp DNA BCT 07-APR-2000
LOCUS Dehalococcoides ethenogenes 16S ribosomal RNA gene, partial
DEFINITION sequence.
ACCESSION AF004928
VERSION AF004928.2 GI:7524025
KEYWORDS
SOURCE Dehalococcoides ethenogenes.
ORGANISM Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Mayo-Gatelli, X., Chien, Y., Gossett, J.M. and Zinder, S.H.
TITLE Isolation of a bacterium that reductively dechlorinates
JOURNAL tetrahydroethene to ethene
MEDLINE Science 276 (5318), 1568-1571 (1997)
97315370
REFERENCE 2 (bases 1 to 1434)
AUTHORS Chien, Y. and Zinder, S.H.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-1997) Microbiology, Cornell University, Wing
Hall, Ithaca, NY 14853, USA
REFERENCE 3 (bases 1 to 1434)
AUTHORS Chien, Y. and Zinder, S.H.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2000) Microbiology, Cornell University, Wing
Hall, Ithaca, NY 14853, USA
REMARK Sequence update by submitter
COMMENT On Apr 7, 2000 this sequence version replaced gi:2232215.
FEATURES
source location/Qualifiers
1. 1434
/organism="Dehalococcoides ethenogenes"
/strain="strain 195"
/db_xref="taxon:61435"
/note="tentative genus and species name; Eubacterium not
closely related to presently known phyla"

rrna <1. .>1434
/product="16S ribosomal RNA"
BASE COUNT 376 a 304 c 440 g 313 t 1 others
ORIGIN
Query Match 94.0%; Score 44.2; DB 1; Length 1434;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggtrtgccgacatagtggttcaytaagccgaaagggcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 TGTGATGGGCTGACATAGTCGTTTCATTAAGCCGCAAGGTGCTTG 186
RESULT 11
AX039540 1443 bp DNA PAT 18-NOV-2000
LOCUS Sequence 7 from Patent WO0063443.
DEFINITION AX039540
ACCESSION AX039540
VERSION AX039540.1 GI:11229562
KEYWORDS
SOURCE Dehalococcoides ethenogenes.
ORGANISM Dehalococcoides ethenogenes
Bacteria; Green non-sulfur bacteria; Dehalococcoides group;
Dehalococcoides.
REFERENCE 1 (bases 1 to 1443)
AUTHORS Hendrickson, E.R. and Ebersole, R.C.
TITLE Nucleic acid fragments for the identification of dechlorinating ba
JOURNAL cteria
PATENT Patent: WO 0063443-A 7 26-OCT-2000;
E.I. DU PONT DE NEMOURS AND COMPANY (US)
FEATURES
source location/Qualifiers
1. 1443
/organism="Dehalococcoides ethenogenes"
/db_xref="taxon:61435"
BASE COUNT 379 a 306 c 443 g 314 t 1 others
ORIGIN
Query Match 94.0%; Score 44.2; DB 6; Length 1443;
Best Local Similarity 85.1%; Pred. No. 5.7e-08;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
QY 1 tggtrtgccgacatagtggttcaytaagccgaaagggcttg 47
||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 140 TGTGATGGGCTGACATAGTCGTTTCATTAAGCCGCAAGGTGCTTG 186
RESULT 12
AC014436 87286 bp DNA HTG 16-NOV-1999
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
DEFINITION pieces.
ACCESSION AC014436
VERSION AC014436.1 GI:6436899
KEYWORDS HTG; HTGS_PHASE2.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 87286)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
REMARK This sequence was identified as CDM:10212830 by the submitter.
COMMENT For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

Location/Qualifiers
1. .87286

BASE COUNT	25135	a	18013	c	18299	g	25839	t	/organism="Drosophila melanogaster" /db_xref="taxon:7227"
ORIGIN									

Query Match	56.28;	Score 26.4;	DB 2;	Length 87286;
Best Local Similarity	71.1%;	Pred. No. 3.5;		
Matches 27; Conservative	5;	Mismatches	6.	Indels 0.

QY 9 gcycacatawgtlyggttcaytaaagccgyaagygtt 46
 ||:||||:||||:||||:||||:||||:||||:||||:|
 Db 53263 GCTTACATATGTTCTTCATTAAGCCTTAAGAACCT 53226

AC010843	170914 bp	DNA

Accession	Definition	LOCUS
AC010843	170914 bp DNA	INV 12-MAY-2001
AC010842	complete sequence.	
AC010843	Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone BACR08A22, complete sequence.	

ACCESSION	AC010843	GI:14028974
VERSION	AC010843.10	
KEYWORDS	HTG.	

Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H., Holt, R.A., Evans, C.A., Gossaye, T.D., Marshall, J.L., and Davis

Sequencing of Drosophila chromosome X, region 12A-12A
unpublished

2 (bases 1 to 170914)

Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Luttenhoff, C., Chamme, M., Chavez, C., Chavakis, J., Chavakis, M.,
Chavakis, N., Chavakis, P., Chavakis, R., Chavakis, S., Chavakis, T.

Direct Submission
Submitted (24-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 54-121, Berkeley, CA 94720, USA
On May 12, 2001 this sequence version replaced gi:13374645.
Sequence submitted by.

new site (<http://www.fruitfly.org/sequence/>) or send email
to fruitfly.berkeley.edu.
Location/Qualifiers
1. 170914
/organism="Drosophila melanogaster"

```

/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="X"
/map="12A-12A"
- /clone="BACR08A22 (D879) "
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6) "
BASE COUNT      50786 a 35003 c 35470 g 49655 t
ORIGIN

```

Query Match	56.2%;	Score 26.4;	DB 3;	Length 170914;
Best Local Similarity	71.1%;	Pred. No. 3.7;		
Matches 27; Conservative	5;	Mismatches	6;	Indels 0.
				Cans 0

QY 9 gcycacatawgtlyggttcaytaaacgcyaaagyctt 46
||:|||||:|:|||||:|||||:||||| |
Db 105904 GCTTACATATGTTGTTTCATTAAACCTTAAGGACCT 105941

Accession	Definition	LOCUS
U01000	Drosophila melanogaster, chromosome X, region 12A-12A, BAC clone BACR30C16, complete sequence.	185810 bp DNA INV 20-MAR-2001

VERSION	AC010995.11	GI:13384330
KEYWORDS	HTG.	
SOURCE	fruit fly.	

Celniker, S.E., Adams, M.D., Krommiller, B., Tyler, D., Holt, R.A., Evans, C.A., Gocayne, J.D., Aamodt, D.C., White, O., Green, E.D., Lasker, J.M., Venter, A., Smith, T.F., Wilson, R., Davis, R.H., and Venter, M.C. 1988. The complete genome sequence of the bacteriophage phiX174. *Nature* 336: 575-579.

Sequencing of Drosophila chromosome X, region 12A-12A
Unpublished
2 (bases 1 to 185810)
Celniker, S.E., Agbayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
Gutenhoff, C., Champe, M., Chavez, C., Chew, M., Christoforoni,
M., D'Amico, A., D'Amico, S.M., and Venter, J.C.

Direct Submission
Submitted (29-SEP-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On Mar 20, 2001 this sequence version replaced gi:6978354.
Sequence submitted by.

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email

mRNA
CDS
gene
gene
CDS
mRNA
gene
CDS

/protein_id="AAF48246.1"
/db_xref="GI:7292854"
/translation="MSEDKMNGRSIRNRLPATIVAILLTFVLVYFLNFHQERPAIY
GMLRSENPSRYNLKMLIAAIAQAQGGLEVLVDVARSROLKERSGKTDEGVNDFTD
ADGRSHCVMKOGLRIEPRVQIFSEEDKEHCKQAHGYDLDPTVLHETAQIPDYVNAQ
DVTVMWDPDLATKEFTEELYEYVTTWCVAAGRPIIGYIHSFENGOTAWAWGNSMS
EYLSNLHPOHSPNNOAPIITVSRSHTAGAKDLARGIFGENVSLTAAGAGYVLQVVA
NNATAYLHRSKIKKWDICAGDAIHALGTMITLNDQLINYGPEESPVNTEGLLTLLE

Search completed: February 9, 2002, 11:31:48
Job time: 2640 sec

Sat Feb 9 11:52:28 2002

us-09-548-998a-8.rge

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 100.64 Seconds
(without alignments)
400.380 Million cell updates/sec

Title: US-09-548-998A-8

Perfect score: 47
Sequence: 1 tgtgttggtggtgacatawt.....Ytaagccggaaggygttg 47

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.1101.*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.2	94.0	47	22	AAC62245 A 16S rDNA sequenc
2	44.2	94.0	1212	22	AAC62239 A 16S rDNA sequenc
3	44.2	94.0	1212	22	AAC62241 A 16S rDNA sequenc
4	44.2	94.0	1212	22	AAC62242 A 16S rDNA sequenc
5	44.2	94.0	1212	22	AAC62243 A 16S rDNA sequenc
6	44.2	94.0	1212	22	AAC62240 A 16S rDNA sequenc
7	44.2	94.0	1212	22	AAC62244 A 16S rDNA sequenc
8	21	44.7	1488	22	AAH74593 A 16S rDNA sequenc
9	21	44.7	81145	22	AAH74593 Nucleotide sequenc
10	20.8	44.3	565	22	AAH06050 Genomic nucleotide
11	20.8	44.3	1560	22	AAF88043 Human cDNA clone (Human Trp2 encodin

C	12	20.8	44.3	2291	18	AAH91957	The novel tyrosina
C	13	20.6	43.8	1491	22	AAH74594	Synthetic nucleoti
C	14	20.6	43.8	3300	22	AAH58569	Human polynucleoti
C	15	20.6	43.8	3316	22	AAH60355	Human polynucleoti
C	16	20.6	43.8	3463	20	AAH76755	Candida RIM1 codin
C	17	20.4	43.4	147	21	AAH69054	Bacteriophage 44AH
C	18	20.4	43.4	478	22	AAH16377	Probe #6310 for ge
C	19	20.4	43.4	787	22	AAH19032	Probe #7718 used t
C	20	20.4	43.4	787	21	AAH74363	Human secreted pro
C	21	20.4	43.4	1464	21	AAH69017	Bacteriophage 44AH
C	22	20.4	43.4	16668	21	AAH69013	Bacteriophage 44AH
C	23	20.2	43.0	1293	22	AAH98420	Human EST-derived
C	24	20.2	43.0	1937	21	AAH47445	Macaque TANGO 232
C	25	20.2	43.0	4594	18	AAH74511	Staphylococcus aur
C	26	20.2	43.0	10207	15	AAH05134	Staphylococcus aur
C	27	20.2	42.6	437	20	AAH90537	EST clone BK498.
C	28	20	42.6	3444	22	AAH83798	Arabidopsis thalia
C	29	19.8	42.1	447	21	AAH44357	Human preproinsuli
C	30	19.8	42.1	480	22	AAH04868	Probe #4859 used t
C	31	19.8	42.1	8355	17	AAH35524	Human interferon-b
C	32	19.6	41.7	375	21	AAH260386	A diacylglycerol a
C	33	19.6	41.7	636	22	AAH03401	Human CDNA clone
C	34	19.6	41.7	689	21	AAH02024	Human colon cancer
C	35	19.6	41.7	1476	22	AAH15252	Human CDNA sequenc
C	36	19.6	41.7	4150	16	AAH082832	Murine stromalin-1
C	37	19.6	41.7	319608	21	AAH51601	Human chromosome 1
C	38	19.6	41.7	319608	22	AAH51601	Human schizopreni
C	39	19.6	41.7	580073	18	AAH58840	Mycoplasma genital
C	40	19.4	41.3	221	20	AAH31851	S. aureus polypept
C	41	19.4	41.3	496	21	AAH15884	Human prostate can
C	42	19.4	41.3	557	21	AAH57352	Pinus radiata tran
C	43	19.4	41.3	558	21	AAH56216	Pinus radiata tran
C	44	19.4	41.3	714	22	AAH67447	C glutamicum codin
C	45	19.4	41.3	782	19	AAH59777	Human secreted pro

ALIGNMENTS

RESULT	ID	Score	Description
1	AAC62245	standard; DNA; 47 BP.	
AC	AAC62245;		
AC	19-MAR-2001	(first entry)	
DE	A 16S rDNA sequence indicative of a chlorinating bacterial strain.		
DE	16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;		
KW	carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;		
KW	trichloroethane; dichloroethylene; chlorinating bacteria; ss.		
OS	Synthetic.		
OS	Dehalococcoides ethenogenes.		
PN	WO200063443-A2.		
PD	26-OCT-2000.		
PF	13-APR-2000; 2000WO-US09883.		
PR	15-APR-1999; 99US-0129511.		
PA	(DUPO) DU PONT DE NEMOURS & CO E I.		
PI	Hendrickson ER, Ebersole RC;		
DR	WPI; 2001-024581/03.		
XX	New 16S rDNA profile derived from Dehalococcoides ethenogenes and		
PT	indicative of a dechlorinating bacterial strain -		
XX			

PS Claim 1; Page 50; 55pp; English.

CC The present sequence represents a consensus 16S rDNA sequence, derived
CC from various strains of Dehalococcoides ethenogenes. The 16S rDNA
CC profile is linked to dechlorinating activity. Bacterial strain
CC comprising the 16S rDNA sequence of the invention are useful for the
CC dechlorination of chlorinated compounds such as carbontetrachloride,
CC tetrachloroethane, chloroform, dichloromethane, trichloroethane,
CC dichloroethylene, vinyl chloride, and chloroaromatics. The 16S rDNA
CC sequence is also useful for identification of new chlorinating bacteria,
CC as well as sub-typing strains of Dehalococcoides ethenogenes.

XX Sequence 47 BP; 9 A; 6 C; 15 G; 10 T; 7 other;

Query Match 94.0%; Score 44.2; DB 22; Length 47;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 47
1 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 47

RESULT 2

AAC62239 standard; DNA; 1212 BP.

AAC62239;

19-MAR-2001 (first entry)

A 16S rDNA sequence indicative of a chlorinating bacterial strain.

16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
trichloroethane; dichloroethylene; chlorinating bacteria; ss.

Dehalococcoides ethenogenes.

WO200063443-A2.

26-OCT-2000.

13-APR-2000; 2000WO-US09883.

15-APR-1999; 99US-0129511.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hendrickson ER, Ebersole RC;

WPI; 2001-024581/03.

New 16S rDNA profile derived from Dehalococcoides ethenogenes and
indicative of a dechlorinating bacterial strain -

Claim 1; Page 47; 55pp; English.

The present sequence represents the 16S rDNA profile of Dehalococcoides
ethenogenes strain PL, isolated from soil surrounding an industrial
site. The 16S rDNA profile is linked to dechlorinating activity.
Bacterial strain comprising the 16S rDNA sequence of the invention are
useful for the dechlorination of chlorinated compounds such as
carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
trichloroethane, dichloroethylene, vinyl chloride, and chloroaromatics.
The 16S rDNA sequence is also useful for identification of new
chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
ethenogenes.

Sequence 1212 BP; 321 A; 252 C; 369 G; 269 T; 1 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 85.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

1 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 47
140 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 186

RESULT 3

AAC62241 standard; DNA; 1212 BP.

AAC62241;

19-MAR-2001 (first entry)

A 16S rDNA sequence indicative of a chlorinating bacterial strain.

16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
trichloroethane; dichloroethylene; chlorinating bacteria; ss.

Dehalococcoides ethenogenes.

WO200063443-A2.

26-OCT-2000.

13-APR-2000; 2000WO-US09883.

15-APR-1999; 99US-0129511.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hendrickson ER, Ebersole RC;

WPI; 2001-024581/03.

New 16S rDNA profile derived from Dehalococcoides ethenogenes and
indicative of a dechlorinating bacterial strain -

Claim 1; Page 48; 55pp; English.

The present sequence represents the 16S rDNA profile of Dehalococcoides
ethenogenes strain DAB, isolated from soil surrounding an industrial
site. The 16S rDNA profile is linked to dechlorinating activity.
Bacterial strain comprising the 16S rDNA sequence of the invention are
useful for the dechlorination of chlorinated compounds such as
carbontetrachloride, tetrachloroethane, chloroform, dichloromethane,
trichloroethane, dichloroethylene, vinyl chloride, and chloroaromatics.
The 16S rDNA sequence is also useful for identification of new
chlorinating bacteria, as well as sub-typing strains of Dehalococcoides
ethenogenes.

Sequence 1212 BP; 318 A; 253 C; 372 G; 269 T; 0 other;

Query Match 94.0%; Score 44.2; DB 22; Length 1212;
Best Local Similarity 85.1%; Pred. No. 2.7e-09;
Matches 40; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

1 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 47
140 tgtgtggcgacatawgttggttcaytaagccggaagggcttg 186

RESULT 4

AAC62242 standard; DNA; 1212 BP.

AAC62242;

XX

ethenogenes strain STF, isolated from soil surrounding an industrial site. The 16S rDNA profile is linked to dechlorinating activity. CC Bacterial strain comprising the 16S rDNA sequence of the invention are CC useful for the dechlorination of chlorinated compounds such as CC carbontetrachloride, tetrachloroethane, chloroform, dichloromethane, CC trichloroethane, dichloroethylene, vinyl chloride, and chloroacromatics. CC The 16S rDNA sequence is also useful for identification of new CC chlorinating bacteria, as well as sub-typing strains of Dehalococcoides CC ethenogenes. CC
Sequence 1335 BP; 348 A; 288 C; 403 G; 296 T; 0 other; CC
XX
SQ

Query Match	94.0%;	Score 44.2;	DB 22;	Length 1335;
Best Local Similarity	85.1%;	Pred. No. 2.8e-09;		
Matches 40;	Conservative	7;	Mismatches 0;	Indels 0;
				Gaps 0;

QY 1 tgtgtrggcygacatawgtlygtgttcaytaagccgyaaggygcttg 47
|||:||||:||||:|:||||:|||||||:||||:|||||
Db 140 tgtgtggcgacataaagtgtgttcactaaagccgtaagtgcttg 186

RESULT	7
AAC62244	
ID	AAC62244 standard; DNA; 1443 BP.

AC AAC62244;

DT 19-MAR-2001 (first entry)

AA A 16S rDNA sequence indicative of a chlorinating bacterial strain.
DE

XX 16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
KW carbon tetrachloride; tetrachloroethane; chloroform; dichloromethane;
KW trichloroethane; dichloroethylene; chlorinating bacteria; ss.

Dehalococcoides ethenogenes.

PN W0200063443-A2.

PD 26-OCT-2000.

13-APR-2000; 2000WO-US09883.

PR 15-APR-1999; 99US-0129511.

(DUPO) DU PONT DE NEMOURS & CO E I.

Hendrickson ER, Ebersole RC; PI

DR WPI; 2001-024581/03.

XX New 16S rDNA profile derived from *Dehalococcoides ethenogenes* and
PT indicative of a dechlorinating bacterial strain -

Example 2; Page 49-50; 55pp; English.

XX The present sequence represents the 16S rDNA profile of *Dehalococcoides*
CC *ethnogenes* strain 195. The 16S rDNA profile is linked to dechlorinating
CC activity. Bacterial strain comprising the 16S rDNA sequence of the
CC invention are useful for the dechlorination of chlorinated compounds such
CC as carbon tetrachloride, tetrachloroethane, chloroform, dichloromethane,
CC trichloroethane, dichloroethylene, vinyl chloride, and chloroaromatics.
CC The 16S rDNA sequence is also useful for identification of new
CC chlorinating bacteria, as well as sub-typing strains of *Dehalococcoides*
CC *ethnogenes*.

XX Sequence 1443 BP; 379 A; 306 C; 443 G; 314 T; 1 other;
SQ

Query Match	94.0%;	Score 44.2;	DB 22;	Length 1443;
Best Local Similarity	85.1%;	Pred. NO. 2.8e-09;		
Matches 40;	Conservative	7;	Mismatches 0;	Indels 0;
				Gaps 0;

QY 1 tgtgtrtggygcycataawgtlygtgttcaytaaagccgcyaaaggygcttg 47
 ||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
 Db 140 tgtgattggygcycgacataaagtgcgttcattaaagccgcaaggtgcttg 186

RESULT	8
AAH74593/c	
ID	AAH74593 standard; DNA; 1488 BP.

AC AAH745937

DT 15-OCT-2001 (first entry)

Nucleotide sequence of a hydantoinase enzyme.

XX methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
KW directed evolution; molecular breeding; hydantoinase; ss.
KW

OS *Pseudomonas putida*.

aa		Location/Qualifiers
FH	Key	1..1488
FT	CDS	/*tag= a
FT		/product= "hydantoinase"
FT		

PN W0200155342-A2.

PD 02-AUG-2001.

31-JAN-2001; 2001WO-US03186.
PF

PR	31-JAN-2000; 2000US-0494921.
PR	08-DEC-2000; 2000US-0734237.

AA
PA (BIOC-) BIOCATALYTICS INC.

PI Rozzell DJ, Bui P, Hua L;

DR WPI; 2001-483235/52.

XX	Designing synthetic nucleic acid sequences for improved amplification,
PT	expression in host cell, by comparing free energy of folding of a
PT	starting polynucleotide and a modified polynucleotide having a codon
PT	replacement -

PS Example 9; page 91-92; 117pp; English.

XX The present sequence encodes a hydantoinase enzyme. The polynucleotide
CC was modified using the method of the invention. The
CC specification describes a method for designing a synthetic
CC polynucleotide. The method comprises providing a starting polynucleotide,
CC determining the predicted free energy of folding per base of the
CC polynucleotide, modifying the polynucleotide by replacing a codon with
CC a different codon to provide a modified polynucleotide, determining free
CC energy of folding per base of the modified polynucleotides, and
CC comparing this with that of the original polynucleotide. The method is
CC useful for developing nucleic acid sequences that enhance expression of
CC the encoded protein in a heterologous host. The design and preparation
CC of the synthetic genes are used in application of gene shuffling,
CC directed evolution and molecular breeding methods. The method allows
CC expression of genes from various organisms such as mammals, plants,
CC yeast, fungi and bacteria in prokaryotic hosts, such as *Escherichia coli*
CC and eukaryotic hosts at commercially viable levels, in particular
CC proteins with low yield such as methionine gamma-lyase from *P. putida*.

Sequence 1488 BP; 285 A; 502 C; 456 G; 245 T; 0 other;

Query Match	44.7%;	Score 21;	DB 22;	Length 1488;
Best Local Similarity	57.4%;	Pred. No. 26;		
Matches 27;	Conservative	5;	Mismatches	15; Indels 0; Gaps 0;

RESULT	9
AAF54868/c	
ID	AAF54868 standard; DNA; 81145 BP.
XX	
AC	AAF54868;

Genomic nucleotide sequence of a human 65h2 protein.

Human; transmembrane protein; 65h2 protein; 593 protein; prostaglandin; thromboxane; KIAA0860 protein; ss.

OS Homo sapiens.

PN .W0200109185-A2.

PD 08-FEB-2001.

28-JUL-2000; 2000WO-US20521.

PR 30-JUL-1999; 99US-0365162.

(MILL-) MILLENNIUM PHARM INC.

Curtis RA;

WPI; 2001-138648/14.

New nucleic acid molecules encoding transmembrane proteins designated 65h2 and 293 are useful for screening assays, detection assays and in predictive medicine -

Disclosure; Fig 1L1-1L36; 215pp; English.

The present sequence encodes a human transmembrane proteins designated 65h2. The specification also describes a transmembrane protein designated 593 and KIA00880. The proteins and polynucleotides can be used for screening assays, detection assays, e.g. chromosome mapping, tissue typing, forensic biology and predictive medicine, e.g. diagnostic assays, prognostic assays, monitoring clinical trials and pharmacogenomics. They polynucleotides can also be used for the recombinant production of 65h2 and 593 proteins, which can be used in assays to detect compounds which modulate their activity. These compounds are useful in the treatment of diseases characterized by aberrant activity or expression of protein which catalyse or facilitate transport of charged organic compound, e.g. prostaglandins and thromboxanes.

Query Match	44.7%;	Score 21;	DB 22;	Length 81145;
Best Local Similarity	67.7%;	Pred. No. 86;		
Matches 21; Conservative	5;	Mismatches	5;	Indels 0;

Y 14 catawgttggttcaaytaagccgyaaggyc 44
||| : ||| : ||| : ||| : ||| : ||
C 58876 CATTTATTGTTCACCTCAAGCCACAAGGTC 58846

Result	10
AAH06050/c	standard; cDNA; 565 bp
AAH06050;	

DT	26-JUN-2001 (first entry)
XX	
DE	Human cDNA clone (5'-primer) SEQ ID NO:2885.
XX	

KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX	
OS	Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 11-JAN-2000; 2000JP-0118776.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 1; SEQ ID 2885; 2537pp + CD ROM; English
 XY

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 565 BP; 134 A; 168 C; 159 G; 99 T; 5 other;

Query Match	44.3%;	Score 20.8;	DB 22;	Length 565;
Best Local Similarity	53.2%;	Pred. No. 23;		
Matches 25; Conservative	7;	Mismatches 15;		

Y
1 tgtgrrtgggcycacatawgtlygttcaytaaacgccyaaggygcttg 47
|||:|||:||||:||||:|||||||: |||:
438 TGGGCTNGAGCAACACATATTGTTCCTTTAAGGCCCTCCAGCTGGTAG 392

RESULT 11
AAFF88043 standard; cDNA; 1560 BP

AC AAF88043;
XX 16-JUL-2001 (first entry)
DT Human trp2 encoding cDNA.
XX
DE
XX Listeria; expression vector; tumor-associated antigen; Trp 1; Trp 2;
KW Melana/MARF-1; cytosolic; attenuated; immunotherapy; malignant melanoma;
KW pigmented tumor; malignant schwannoma; vaccination; tyrosinase;
KW antigen-presenting cell; ss.
XX
OS Homo sapiens.
XX WO200127295-A1.
PN
XX 19-APR-2001.
PD
XX 13-OCT-2000; 2000WO-DE03629.
PF
XX 14-OCT-1999; 99DE-1049594.
XX
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
PA
XX Schendendorf D, Paschen A, Chakraborty T, Domann E;
PI
XX WPI: 2001-282041/29.
DR P-PSDB; AAB86041.
DR
XX Listeria expression vector for immunotherapy, particularly of malignant
PT melanoma, comprises a DNA sequence encoding tumor-associated antigens -
PS Disclosure; Fig 3; 41pp; German.
XX
XX This invention describes a novel Listeria expression vector (A) for
CC immunotherapy which comprises a promoter (P), functional in Listeria,
CC operably linked to a DNA sequence (I) encoding one of the
CC tumor-associated antigens (II) human tyrosinase, Trp 1 or 2, or
CC Melana/MARF-1. The products of the invention have cytostatic activity.
CC Recombinant attenuated Listeria containing (A) are useful for
CC immunotherapy (prophylactic, adjuvant or therapeutic), specifically of
CC malignant melanoma (but also other pigmented tumors such as malignant
CC schwannoma), particularly as a replacement for radiotherapy. Using
CC attenuated Listeria as carrier for (A) provides a simple way of
CC vaccination, since antigen-presenting cells acquire tumor-associated
CC antigens by natural infection, eliminating the need for labor-intensive
CC ex vivo modification of autologous cells. This sequence encodes the
CC human Trp2 protein described in the method of the invention.
XX
SQ Sequence 1560 BP; 388 A; 384 C; 387 G; 401 T; 0 other;
QY
Query Match 44.3%; Score 20.8; DB 22; Length 1560;
Best Local Similarity 59.5%; Pred. No. 32;
Matches 25; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
DB 2 gtgttgagcgcacatawgttggtcaytaagccgcyaggyg 43
1093 gtgatgagccttcataattgtgtcattcctcctgaacggyg 1134
RESULT 12
ID AAT91957 standard; cDNA; 2291 BP.
XX
AC AAT91957;
XX
DT 20-MAR-1998 (first entry)
XX
DE The novel tyrosinase-related protein 2 (TRP-2) gene.
XX
XX Tyrosinase related protein 1 gene; TRP-1; TRP-2; gp75; tumour antigen;
KW tumour infiltrating lymphocyte; TIL; TIL586; cancer peptide; TRP-2;
KW alternative reading frame; cancer detection; pre-cancer detection;

KW melanoma; ds.
XX
XX Homo sapiens.
OS
XX
XX Key
FT CDS 415..1974
FT Location/Qualifiers
/*tag= a
WO9729195-A2.
XX
XX 14-AUG-1997.
PD
XX
XX 06-FEB-1997; 97WO-US02186.
PF
XX 04-OCT-1996; 96US-0725736.
PR 09-FEB-1996; 96US-0599602.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Rosenberg SA, Wang R;
PI
XX WPI: 1997-415349/38.
DR P-PSDB; AAW30826.
DR
XX
XX Cancer antigen peptide(s) derived from the tyrosinase-related
PT protein 1 or 2 - useful for detecting, preventing or treating a
PT cancer in a mammal, especially melanoma
XX
XX
PS Claim 13; Pages 79-81; 11pp; English.
XX
XX The present sequence represents part of the novel tyrosinase related
CC protein 2 (TRP-2) gene. This gene encodes tumour antigens recognised
CC by tumour infiltrating lymphocyte (TIL) 586. Novel cancer peptides have
CC also been identified in TRP-1. The peptides are recognised by a major
CC histocompatibility complex (MHC) class I T-lymphocyte. The nucleic acids
CC encoding the cancer peptides or TRP-2 can be used to detect a cancer or
CC pre-cancer in a mammal, especially by detecting the presence of the
CC alternative ORF 3 of the TRP-1 gene or the sequence encoding the novel
CC tumour antigen TRP-2. Vectors and recombinant viruses containing
CC antigen peptide encoding nucleic acids, antibodies raised against the
CC peptides, or the peptides themselves can be used to prevent or treat
CC a cancer in a mammal, especially a melanoma.
XX
SQ Sequence 2291 BP; 641 A; 510 C; 545 G; 595 T; 0 other;
QY
Query Match 44.3%; Score 20.8; DB 18; Length 2291;
Best Local Similarity 59.5%; Pred. No. 36;
Matches 25; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
DB 2 gtgttgagcgcacatawgttggtcaytaagccgcyaggyg 43
1507 gtgatgagccttcataattgtgtcattcctcctgaacggyg 1548
RESULT 13
ID AAH74594/C standard; DNA; 1491 BP.
XX
AC AAH74594;
XX
DT 15-OCT-2001 (first entry)
XX
DE Synthetic nucleotide sequence of a hydantoinase enzyme.
XX
XX Methionine gamma-lyase; mdeA gene; free folding energy; gene shuffling;
KW directed evolution; molecular breeding; hydantoinase; ss.
XX
XX Synthetic.
OS Pseudomonas putida.
OS
XX
XX Key
FT CDS Location/Qualifiers
1..1491

/tag= a
/product= "hydantoinase"

WO200155342-A2.

02-AUG-2001.

31-JAN-2001; 2001WO-US03186.

31-JAN-2000; 2000US-0494921.
08-DEC-2000; 2000US-0734237.

(BIOC-) BIOCATALYTICS INC.

Rozzell DJ, Bul P, Hua L;

WPI; 2001-483235/52.
P-PSDB; AAG63576.

Designing synthetic nucleic acid sequences for improved amplification, expression in host cell, by comparing free energy of folding of a starting polynucleotide and a modified polynucleotide having a codon replacement

Claim 9; Page 94; 11'pp; English.

The present sequence encodes a modified hydantoinase. The polynucleotide was modified using the method of the invention. The polynucleotide specification describes a method for designing a synthetic polynucleotide. The method comprises providing a starting polynucleotide, determining the predicted free energy of folding per base of the polynucleotide, modifying the polynucleotide by replacing a codon with a different codon to provide a modified polynucleotide, determining free energy of folding per base of the modified polynucleotide, and comparing this with that of the original polynucleotide. The method is useful for developing nucleic acid sequences that enhance expression of the encoded protein in a heterologous host. The design and preparation of the synthetic genes are used in application of gene shuffling, directed evolution and molecular breeding methods. The method allows expression of genes from various organisms such as mammals, plants, yeast, fungi and bacteria in prokaryotic hosts, such as *Escherichia coli* and eukaryotic hosts at commercially viable levels, in particular proteins with low yield such as methionine gamma-lyase from *P. putida*.

Sequence 1491 BP; 250 A; 403 C; 412 G; 426 T; 0 other;

Query Match 43.8%; Score 20.6; DB 22; Length 1491;
Best Local Similarity 35.3%; Pred. No. 38;
Matches 26; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

OY 1 tgtgtggcgyacatawgttggttcaytaagccgyaaggygcttg 47
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 496 TGTGTCAGCCGCCATATGCGTTTGTGTAAGCCATGAAGTGTG 450

RESULT 14

AAI58569
ID AAI58569 standard; cDNA; 3300 BP.

AAI58569;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 772.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

OS Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

25-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

14-SEP-2000; 2000US-0662191.

19-OCT-2000; 2000US-0693036.

29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;

WPI; 2001-442253/47.
P-PSDB; AAM39413.

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries

Claim 1; SEQ ID NO 772; 10078pp; English.

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localized neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 3300 BP; 1025 A; 593 C; 695 G; 987 T; 0 other;

Query Match 43.8%; Score 20.6; DB 22; Length 3300;
Best Local Similarity 62.2%; Pred. No. 49;
Matches 23; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 10 cygacatawgttggttcaytaagccgyaaggygctt 46
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1520 ctgcatatttaagttcattcaagcagcaaatcctt 1556

RESULT 15

AAI60355
ID AAI60355 standard; cDNA; 3316 BP.

AAI60355;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 4344.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

XX 09-JUL-2000; 2000US-0598042.

XX 19-JUL-2000; 2000US-0620312.

XX 03-AUG-2000; 2000US-0653450.

XX 14-SEP-2000; 2000US-0662191.

XX 19-OCT-2000; 2000US-0693036.

XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM41199.

XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -

XX Claim 1; SEQ ID NO 4344; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed
XX specification.

XX Sequence 3316 BP; 1013 A; 595 C; 720 G; 983 T; 5 other;

XX SQ

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Search completed: February 9, 2002, 11:33:39
Job time: 2751 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 38.58 Seconds
(without alignments)
275.906 Million cell updates/sec

Title: US-09-548-998A-8

Perfect score: 47

Sequence: 1 tgtgtgtgcygacatawgt.....ytaaagccgyaaggycttg 47

Scoring table: IDENTITY..NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents..NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_5/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20.8	44.3	2291	2	US-08-725-736D-1	Sequence 1, Appli
2	20.8	44.3	2291	3	US-09-162-368B-1	Sequence 1, Appli
3	20.8	44.3	2291	3	US-09-161-877B-1	Sequence 1, Appli
4	20.6	43.8	3463	4	US-09-189-462-3	Sequence 3, Appli
5	20.2	43.0	10207	1	US-08-920-812-2	Sequence 2, Appli
6	20.2	43.0	10207	1	US-08-920-827-2	Sequence 2, Appli
7	20.2	43.0	10207	1	US-08-921-177-2	Sequence 2, Appli
8	20.2	43.0	10207	1	US-08-920-828-2	Sequence 2, Appli
9	20.2	43.0	10207	2	US-08-920-828-2	Sequence 2, Appli
10	20.2	43.0	10207	2	US-08-920-828-2	Sequence 2, Appli
11	19.8	42.1	8355	4	US-08-406-030A-23	Sequence 144, App
12	19.4	41.3	2499	4	US-09-205-283-1	Sequence 23, Appli
13	19.4	41.3	2499	4	US-09-205-283-1	Sequence 1, Appli
14	19.4	41.3	2520	3	US-08-968-563-10	Sequence 11, Appli
15	19.4	41.3	2520	3	US-08-969-683A-10	Sequence 10, Appli
16	19.4	41.3	19227	3	US-09-090-793-13	Sequence 13, Appli
17	19.4	41.3	40138	3	US-09-090-793-13	Sequence 12, Appli
18	19.2	40.9	589	4	US-09-385-982-454	Sequence 454, App
19	19.2	40.9	1764	4	US-08-980-832-36	Sequence 36, Appli
20	19.2	40.9	2023	2	US-08-417-210A-136	Sequence 136, App
21	19.2	40.9	2023	2	US-08-417-210A-139	Sequence 139, App
22	19.2	40.9	2060	2	US-08-417-210A-142	Sequence 142, App
23	19.2	40.9	3040	1	US-08-446-794A-1	Sequence 1, Appli
24	19.2	40.9	3040	1	US-08-750-007-2	Sequence 2, Appli
25	19.2	40.9	3040	2	US-08-945-024-1	Sequence 1, Appli
26	19.2	40.9	3808	2	US-08-417-210A-79	Sequence 79, Appli
27	19.2	40.9	15328	2	US-08-888-497-33	Sequence 33, Appli

28	19.2	40.9	15328	5	PCT-US94-07926-33	Sequence 33, Appli
29	19	40.4	1688	2	US-08-793-410-1	Sequence 1, Appli
30	19	40.4	1688	2	US-08-793-410-27	Sequence 27, Appli
31	19	40.4	3515	1	US-08-596-985-1	Sequence 1, Appli
32	19	40.4	7333	4	US-08-766-528-2	Sequence 2, Appli
33	19	40.4	8060	4	US-08-766-528-1	Sequence 1, Appli
34	19	40.4	8132	4	US-08-766-528-3	Sequence 3, Appli
35	19	40.4	10136	5	PCT-US95-16216-2	Sequence 2, Appli
36	19	40.0	387	1	US-08-423-383-2	Sequence 2, Appli
37	18.8	40.0	1164	6	5240849-1	Sequence 2, Appli
38	18.8	40.0	1164	6	5240849-1	Sequence 2, Appli
39	18.8	40.0	2749	6	5240849-1	Sequence 2, Appli
40	18.8	40.0	1318	4	US-08-986-304-1	Sequence 1, Appli
41	18.6	39.6	3240	1	US-08-294-189-2	Sequence 2, Appli
42	18.6	39.6	5494	4	US-08-910-864-10	Sequence 10, Appli
43	18.6	39.6	5510	1	US-08-123-161A-7	Sequence 7, Appli
44	18.6	39.6	5510	1	US-08-123-161A-7	Sequence 7, Appli
45	18.6	39.6	5510	1	US-08-483-278-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-725-736D-1
; Sequence 1, Application US/08725736D
; Patent No. 5831016
; GENERAL INFORMATION:
; APPLICANT: WANG, R.F.; ROSENBERG, S. A.
; TITLE OF INVENTION: IDENTIFICATION OF TRP-2 AS
; TITLE OF INVENTION: A HUMAN TUMOR ANTIGEN RECOGNIZED BY CYTOTOXIC T
; NUMBER OF INVENTION: LYMPHOCYTES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/725,736D
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435.
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,602
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2291
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; MOLECULE TYPE:
; DESCRIPTION: CDNA
; FEATURE:
; NAME/KEY: TRP-2 GENE
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:

LOCATION: 1
IDENTIFICATION METHOD: 2
OTHER INFORMATION: 3
PUBLICATION INFORMATION: 4
AUTHORS: YOKAYAMA, ET AL. 5
TITLE: MOLECULAR CLONING AND FUNCTIONAL 6
TITLE: ANALYSIS OF A CDNA CODING FOR HUMAN 7
TITLE: DOPACHROME TAUTOMERASE/TYROSINASE-RELATED 8
TITLE: PROTEIN-2. 9
JOURNAL: BIOCHIM. BIOPHYS. ACTA. 10
VOLUME: 1217 11
ISSUE: 12
PAGES: 317-321 13
100

Query Match	Score	DB 3	Length
44.38;	20.8;	2291;	
50.59;	7.8;		

QY 2 gtgrrggcgacatawgttggttcayttaagccgyaagyyg 43

RESULT 3

US-09-161-877B-1
; Sequence 1, Application US/09161877E

GENERAL INFORMATION:
APPLICANT: WANG, R.F.; ROSENBERG, S.A.
TITLE OF INVENTION: IDENTIFICATION OF TRP-2
TITLE OF INVENTION: AS A HUMAN TUMOR ANTIGEN RECOGNIZED BY
TITLE OF INVENTION: CYTOTOXIC T LYMPHOCYTES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK

```

; ;
; ; COMPUTER READABLE FORM:
; ; MEDIUM TYPE: FLOPPY DISK
; ; COMPUTER: IBM PC COMPATIBLE
; ; OPERATING SYSTEM: PC-DOS/MS-DOS
; ; SOFTWARE: MICROSOFT WORD 97
; ; CURRENT APPLICATION DATA:
; ; MS/09/161 8

```

CONVENT APPLICATION NUMBER: US/09/161,877B
 APPLICATION NUMBER: US/09/161,877B
 FILING DATE: 28-SEPT-1998
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/725,736
 FILING DATE: 04-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/599,602
 00 SEP-1906

ATTORNEY/AGENT INFORMATION:
NAME: KATHRYN M. BROWN
REGISTRATION NUMBER: 34,556
REFERENCE/DOCKET NUMBER: 2026-4243US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

```

1 INFORMATION FOR SEQ ID NO:
2
3 SEQUENCE CHARACTERISTICS
4
5 LENGTH: 2291
6
7 TYPE: NUCLEIC ACID
8
9 STRANDEDNESS: DOUBLE
10
11 TOPOLOGY: UNKNOWN
12
13 MOLECULE TYPE:

```

```

; INFORMATION FOR SEQ ID NO: 1:
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 2291
;   TYPE: NUCLEIC ACID
;   STRANDEDNESS: DOUBLE
;   TOPOLOGY: UNKNOWN
;
; MOLECULE TYPE:
;
;

```

```

1  DESCRIPTION:  CDNA
2  FEATURE:
3  NAME/KEY:  TRP-2 GENE
4  LOCATION:
5  IDENTIFICATION METHOD:
6  OTHER INFORMATION:
7  PUBLICATION INFORMATION:
8  AUTHORS:  YOKAYAMA, ET AL.
9  TITLE:  MOLECULAR CLONING AND FUNCTIONAL
10 TITLE:  ANALYSTS OF A CDNA CODING FOR HUMAN
11 TITLE:  DOPACHROME TAUTOMERASE/TYROSINASE
12 TITLE:  RELATED PROTEIN-2.
13 JOURNAL:  BIOCHIM. BIOPHYS. ACTA.
14 VOLUME:  1217
15 ISSUE:
16 PAGES:  317-321
17 DATE:  . 1994
18
19 US-09-161-877B-1

```

Query Match	44.38;	Score 20.8;	DB 3;	Length 2291;
Best Local Similarity	59.58;	Pred. No. 7.8;		
Matches 25;	Conservative 5;	Mismatches 12;	Indels 0;	Gaps 0;

DY 2 gtgtctgagcgcacatawtltygttcaytaaagccgyaagy 43
|||::|||:: |||:: |:|||||::| || ||| ||
Db 1507 GTGATGAGCCTTCATAATTGGTTCATCCCTTCGTGACGG 1548

```

RESULT 4
US-09-189-462-3/c
; Sequence 3, Application US/09189462
; Patent No. 6303302
; GENERAL INFORMATION:
; APPLICANT: Rupp, Steffan
; APPLICANT: Roberson, Laura
; APPLICANT: Summers, Eric F.
; APPLICANT: Hecht, Peter
; APPLICANT: Roberts, Radclyffe
; APPLICANT: Madhani, Hiran
; APPLICANT: Styles, Cora Ann
; APPLICANT: Lo, Hsiu-Jung
; APPLICANT: Sherman, Amir
; APPLICANT: Cali, Brian
; APPLICANT: Fink, Gerald
; TITLE OF INVENTION: REGULATION OF FUNGAL GENE EXPRESSION
; FILE REFERENCE: 50078/003002
; CURRENT APPLICATION NUMBER: US/09/189,462
; CURRENT FILING DATE: 1998-11-10
; EARLIER APPLICATION NUMBER: 60/066,129
; EARLIER FILING DATE: 1997-11-19
; EARLIER APPLICATION NUMBER: 60/066,308
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: 60/066,462
; EARLIER FILING DATE: 1997-11-24
; EARLIER APPLICATION NUMBER: 60/078,610
; EARLIER FILING DATE: 1998-03-19
; EARLIER APPLICATION NUMBER: 60/094,523
; EARLIER FILING DATE: 1998-07-29
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3463
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-189-462-3

```

Query Match	43.88;	Score 20.6;	DB 4;	Length 3463;
Best Local Similarity	55.3%;	Pred. No. 11;		
Matches 26;	Conservative 6;	Mismatches 15;	Indels 0;	Gaps 0;

```

Db      787  TGTGGTTGGTAGCCATAAGTTGGTGGTGTAAATAATAACATGGTTG  741
          ||| : ||| | |||| : |||| : ||| : |||

```

```

RESULT      5
US-08-920-812-2/c
; Sequence 2, Application US/08920812
; Patent No. 5763188
;
; GENERAL INFORMATION:
;   APPLICANT: Ohno, Tsuneya
;   APPLICANT: Matsuhisa, Akio
;   APPLICANT: Uehara, Hirotsugu
;   APPLICANT: Eda, Soji
;   TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
;   NUMBER OF SEQUENCES: 25
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;     STREET: 6300 Sears Tower, 233 South Wacker Drive
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: United States of America
;     ZIP: 60606-6402
;
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/920,812
;     FILING DATE: 29-AUG-1997
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/362,577
;     FILING DATE: 27-MAR-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Rln-laures, Li-Hsien
;     REGISTRATION NUMBER: 33,547
;     REFERENCE/DOCKET NUMBER: 19036/32420
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312/474-6300
;     TELEFAX: 312/474-0448
;     TELEX: 25-3856
;
;   INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 10207 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: linear
;     MOLECULE TYPE: Genomic DNA
;     ORIGINAL SOURCE:
;       ORGANISM: Staphylococcus aureus
;       STRAIN: Clinical Isolate SA-24
;
US-08-920-812-2

```

Query Match	43.0%;	Score 20.2;	DB 1;	Length 10207;
Best Local Similarity	59.5%;	Pred. No. 23;		
Matches	22;	Conservative 6;	Mismatches 9;	Indels 0;
Gaps				0;
QY	10	CYACATAAGTGGTCATCAAGCGYAAGYGCT	46	
		: : : : : : : : : : : : : : : :		
Db	6253	CTGACATAACACTTTCCACAATACGCCGAAGGTACTT	6217	

```

RESULT      6
US-08-920-827-2/c
; Sequence 2, Application US/08920827
; Patent No. 5770375
;
; GENERAL INFORMATION:
;
; APPLICANT:      Ohno, Tsuneya
;
; APPLICANT:      Matsuhisa, Akio
;
; APPLICANT:      Uehara, Hirotsugu
;
; APPLICANT:      Eda, Soji

```

TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,827
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-920-827-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 10 cygacatawgttyggttcaytaagccgyaagygtt 46
1:|||||: : ||||: | | ||: |||
Db 6253 CTGACATAACATCTTTCACATACGCGCAAGGTACTT 6217

RESULT 7
US-08-921-177-2/c
; Sequence 2, Application US/08921177
; Patent No. 5798211
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921,177
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 10207 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus aureus
STRAIN: Clinical Isolate SA-24
US-08-921-177-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 10 cygacatawgttyggttcaytaagccgyaagygtt 46
1:|||||: : ||||: | | ||: |||
Db 6253 CTGACATAACATCTTTCACATACGCGCAAGGTACTT 6217

RESULT 8
US-08-362-577C-2/c
; Sequence 2, Application US/08362577C
; Patent No. 5807673
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Eda, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,577C
; FILING DATE: 27-MAR-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

;; INFORMATION FOR SEQ ID NO: 2;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10207 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: Genomic DNA
;; ORIGINAL SOURCE:
;; ORGANISM: Staphylococcus aureus
;; STRAIN: Clinical Isolate SA-24
US-08-362-577C-2

Query Match 43.0%; Score 20.2; DB 1; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 10 cygacatawgtlygttcaytaagccgyaagygtt 46
|:|||||: : |||: | | :|||: |||
Db 6253 CTGACATACATCTTTCACATACGCGCAAGTACTT 6217

RESULT 9

US-08-920-828-2/c
; Sequence 2, Application US/08920828
; Patent No. 5853998

GENERAL INFORMATION:

;; APPLICANT: Ohno, Tsuneya
;; APPLICANT: Matsuhisa, Akio
;; APPLICANT: Uehara, Hirotsugu
;; APPLICANT: Eda, Soji
;; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
;; NUMBER OF SEQUENCES: 25
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;; STREET: 6300 Sears Tower, 233 South Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: United States of America
;; ZIP: 60606-6402

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/920,828
;; FILING DATE: 29-AUG-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/362,577
;; FILING DATE: 27-MAR-1995

;; ATTORNEY/AGENT INFORMATION:
;; NAME: Rin-Laures, Li-Hsien
;; REGISTRATION NUMBER: 33,547
;; REFERENCE/DOCKET NUMBER: 19036/32420

;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312/474-6300
;; TELEFAX: 312/474-0448
;; TELEX: 25-3856

;; INFORMATION FOR SEQ ID NO: 2;
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10207 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: Genomic DNA
;; ORIGINAL SOURCE:
;; ORGANISM: Staphylococcus aureus
;; STRAIN: Clinical Isolate SA-24

US-08-920-828-2

Query Match 43.0%; Score 20.2; DB 2; Length 10207;
Best Local Similarity 59.5%; Pred. No. 23;
Matches 22; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 10 cygacatawgtlygttcaytaagccgyaagygtt 46
|:|||||: : |||: | | :|||: |||
Db 6253 CTGACATACATCTTTCACATACGCGCAAGTACTT 6217

RESULT 10

US-09-060-756-144/c
; Sequence 144, Application US/09060756
; Patent No. 6183957

GENERAL INFORMATION:

;; APPLICANT: Cole, Stewart
;; APPLICANT: Buchrieser-Brosch, Roland
;; APPLICANT: Gordon, Stephen
;; APPLICANT: Billault, Alain
;; TITLE OF INVENTION: METHOD FOR ISOLATING A POLYNUCLEOTIDE OF INTEREST FROM
;; TITLE OF INVENTION: THE GENOME OF A MYCOBACTERIUM USING A BAC-BASED DNA
;; FILE REFERENCE: 3495-0169
;; CURRENT APPLICATION NUMBER: US/09/060,756
;; CURRENT FILING DATE: 1998-04-16
;; NUMBER OF SEQ ID NOS: 743
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 144
;; LENGTH: 269
;; TYPE: DNA
;; ORGANISM: Mycobacterium tuberculosis

;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (various positions within the sequence)
;; OTHER INFORMATION: applicants are uncertain of bases designated as "n"

US-09-060-756-144

Query Match 42.6%; Score 20; DB 4; Length 269;
Best Local Similarity 59.1%; Pred. No. 8.7;
Matches 26; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 4 grtggcygacatawgtlygttcaytaagccgyaagygtt 47
|:| | | | | : | | | | | : |||: |||: |||: |||
Db 243 GGTGCCCGGACAGGAAAGCGTTCAATTGCGCGCAAGTGCTTG 200

RESULT 11

US-08-406-030A-23
; Sequence 23, Application US/08406030A
; Patent No. 6270989

GENERAL INFORMATION:

;; APPLICANT: Treco, Douglas A.
;; APPLICANT: Heartlein, Michael W.
;; APPLICANT: Hauge, Brian M.
;; APPLICANT: Selden, Richard F
;; TITLE OF INVENTION: Protein Production and Delivery
;; NUMBER OF SEQUENCES: 30
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
;; STREET: Two Militia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02173

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/406,030A
;; FILING DATE: 17-MAR-1995
;; CLASSIFICATION: 435

Query Match	41.38;	Score 19.4;	DB 4;	Length 2499;
Best Local Similarity	53.58;	Pred. No. 32;		

RESULT 14
US-08-968-563-10/c
; Sequence 10, Application US/08968563
; Patent No. 6013494
; GENERAL INFORMATION:
; APPLICANT: CHARLES E. NAKAMURA
; APPLICANT: ANTHONY A. GATENBY
; APPLICANT: AMY (KUANG-HUA) HSU
; APPLICANT: RICHARD D. LA REAU
; APPLICANT: SHARON L. HAYNIE
; APPLICANT: MARIA DIAZ-TORRES
; APPLICANT: DONALD E. TRIMBUR
; APPLICANT: GREGORY M. WHITED
; APPLICANT: VASANTHA NAGARAJAN
; APPLICANT: MARK S. PAYNE
; APPLICANT: STEPHEN K. PICATAGGIO
; APPLICANT: RAMESCH V. NAIR
; TITLE OF INVENTION: METHOD FOR THE RECOMBINANT
; TITLE OF INVENTION: PRODUCTION OF 1,3-PROPANEDIOL
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898

Sat Feb 9 11:52:29 2002

us-09-548-998a-8.rni

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2002, 10:47:48 ; Search time 1178.28 Seconds

(without alignments)
428.635 Million cell updates/sec

Title: US-09-548-998a-8

Perfect score: 47

Sequence: 1 tgtgtgtgcygacatawgt.....ytaaagccgyaaggygcttg 47

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_est1:*
11: gb_est2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	24.8	52.8	869	AZ136776	AZ136776 SP_0175_A
C 2	24.4	51.9	510	AO872190	AO872190 nbe0047K
C 3	24	51.1	438	AI777700	AI777700 EST258495
C 4	23.8	50.6	564	BI181099	BI181099 UNL-P-FN-
C 5	23.8	50.6	564	BI186129	BI186129 UNL-P-FN-
C 6	23.2	49.4	1028	CNS02BOM	AL190111 Tetraodon
C 7	23	48.9	469	AQ761792	AQ761792 HS_3147_A
C 8	23	48.9	689	AQ312407	AQ312407 RPT11-10
C 9	23	48.9	835	BG585334	BG585334 EST487098
C 10	22.8	48.5	267	TA358E08P	AL494110 T. brucei
C 11	22.6	48.1	520	AA935599	AA935599 op26f06.s
C 12	22.6	48.1	565	BG948475	BG948475 IPL_10_B1

C 13	22.6	48.1	598	AQ699189	AO699189 HS_5565_A
C 14	22.6	48.1	1008	BG331740	BG331740 602433161
C 15	22.4	47.7	354	BG089520	BG089520 ut68h03.y
C 16	22.4	47.7	422	AM124911	AM124911 UI-M-BH2
C 17	22.4	47.7	502	BE358808	BE358808 DGL_32_G1
C 18	22.4	47.7	525	TA288D10P	AL487170 T. brucei
C 19	22.4	47.7	575	BI183356	BI183356 UNL-P-FN-
C 20	22.4	47.7	683	BE650336	BE650336 UI-M-BH2
C 21	22.4	47.7	905	BG398645	BG398645 602440250
C 22	22.2	47.2	403	AV630336	AV630336 AV630336
C 23	22.2	47.2	430	AV640450	AV640450 OSJNB007
C 24	22.2	47.2	449	AV640450	AV640450 nbe0018J
C 25	22.2	47.2	472	AO868010	AO868010 PS20f09.y
C 26	22.2	47.2	489	AO500765	AO500765 V51F10.mt
C 27	22.2	47.2	532	AM963007	AM963007 EST375080
C 28	22.2	47.2	659	AM963007	AM963007 PS20f09.y
C 29	22.2	46.8	613	BG353081	BG353081 PS20f09.y
C 30	22	46.8	642	AZ730134	AZ730134 RPT-24-1
C 31	22	46.8	673	BG732137	BG732137 PS20f09.y
C 32	22	46.8	1006	CNS060VY9	AL417751 T7 end of
C 33	21.8	46.4	445	AA264746	AA264746 LD08428.5
C 34	21.8	46.4	457	AA951187	AA951187 LD31527.5
C 35	21.8	46.4	463	AA441387	AA441387 LD16303.5
C 36	21.8	46.4	532	AM531096	AM531096 UI-R-C4-a
C 37	21.8	46.4	556	BI162978	BI162978 RE02264.5
C 38	21.8	46.4	578	AZ102232	AZ102232 RPT-23-4
C 39	21.8	46.4	581	AZ824456	AZ824456 2M0099J06
C 40	21.8	46.4	586	TA388A06P	AL498758 T. brucei
C 41	21.8	46.4	835	BG771676	BG771676 602720127
C 42	21.8	46.4	870	BF537581	BF537581 602053346
C 43	21.8	46.4	903	CNS06Y2P	AL420503 T3 end of
C 44	21.8	46.4	942	BF793898	BF793898 602254501
C 45	21.8	46.4	952	BE886937	BE886937 601506418

ALIGNMENTS

RESULT 1
AZ136776/c
LOCUS
DEFINITION
SP_0175_AL_F01_T7A Strongylocentrotus purpuratus, purple sea urchin
'sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone Plate=175 Col=1 Row=K, DNA sequence.

ACCESSION
AZ136776
VERSION
AZ136776.1 GI:8288679
KEYWORDS
GSS.
SOURCE
Strongylocentrotus purpuratus.
ORGANISM
Strongylocentrotus purpuratus.
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.

REFERENCE
1 (bases 1 to 869)
Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P., Biondi, T.R.,
Swartzell, S., Wallace, J.C., Poustka, A.J., Livingston, B.T., Wray,
G.A., Ettensohn, C.A., Lehrach, H., Britten, R.J., Davidson, E.H. and
Hood, L.

TITLE
A sea urchin genome project: Sequence scan, virtual map, and
additional resources
JOURNAL
Proc. Natl. Acad. Sci. U. S. A. 97 (17), 9514-9518 (2000)
MEDLINE
20402566
COMMENT
Contact: Cameron, RA, Davidson, EH, Hood, L
Division of Biology 156-29
California Institute of Technology
Pasadena California 91125, USA
Tel: (626) 395-8421
Fax: (626) 793-3047
Email: acameron@caltech.edu
Plate: 175 row: K column: 1
Seq primer: T7
Class: BAC ends
High quality sequence stop: 869.

FEATURES

Location/Qualifiers

```

source
1. 869
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="Plate=175 Col=1 Row=K"
/clone_lib="Strongylocentrotus purpuratus, purple sea
urchin, sperm genomic BAC library"
/note="Organ: sperm; Vector: BACE3.6; BAC Clones in E-Coli
DH10B"

BASE COUNT      176 a      160 c      207 g      326 t

ORIGIN

Query Match      52.8%; Score 24.8; DB 13; Length 869;
Best Local Similarity 61.9%; Pred. NO. 25;

Matches 26; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY      1 tgtgtgtgagcyacatawgtlygttcayttaagccgyaaggy 42
      |||:|||||:| ||| :| ||| ||||| ||| |||:
DB      239 TGAGATGGGCTGCCATCAGACGGCTCACCAAGTAGCGAGGC 198

```

RESULT 2
AO872190/c
LOCUS
DEFINITION
nbeb0047K03f CUGI Rice BAC Library (EccorI) Oryza sativa genomic
clone nbeb0047K03f, DNA sequence.
GSS 03-NOV-1999
AO872190 510 bp DNA
AO872190
AO872190.1 GI:6222641
GSS.
Oryza sativa.
Oryza sativa
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 510)
Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA

FEATURES

```

/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbeb0047K03f"
/clone_lib="CUGI Rice BAC Library (EcoRI)"
/tissue_type="leaf"
/lab_host="E. coli DH10B"
/notes="Vector: pBACindigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
haploid genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa.
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,296 clones with an average insert size

```

of 121 Kb providing approximately 15 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Three high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening and can be requested from the Clemson University BAC/EST Resource Center (www.genome.clemson.edu). "

Query Match	51.9%;	Score 24.4;	DB 13;	Length 510;
Best Local Similarity	60.9%;	Pred. No. 32;		
Matches	28;	Conservative	6;	Mismatches 12;
				Indels 0;
				Gaps 0;
OY	1	tgtgtrggcgcygacatawgttygttcaytaagccgyaaggygctt	46	
Db	420	TGTAATGAGCCTTTCTCATGTGATTTACTTAAGCTGTAAAGCGCGGTT	375	

RESULT	3	EST	18-MAY-2001
LOCUS	AI7777700	438 bp	MRNA
DEFINITION	EST258495 tomato susceptible, Cornell Lycopersicon esculentum clone CLES2F21, mRNA sequence.		
ACCESSION	AI7777700		
VERSION	AI7777700.1	GI:5275657	
KEYWORDS	EST.		
SOURCE	tomato.		
ORGANISM	Lycopersicon esculentum		

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Lycopersicon esculentum					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;					
Lycopersicon.					
1 (bases 1 to 438)					
D', Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,					
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman					
, C.L., Niernan, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,					
Giovannoni, J.J. and Martin, G.B.					
Generation of ESTs from Pseudomonas susceptible tomato					
Unpublished (1999)					
Contact: CUGI					

TITLE	JOURNAL	COMMENT	FEATURES	source
Generation of ESTs from <i>Pseudomonas</i> susceptible	Unpublished (1999)	Contact: CUGI	5 prime sequence.	1. .438
		Clemson University Genomics Institute	Location/Qualifiers	
		Clemson University		
		100 Jordan Hall, Clemson, SC 29634, USA		
		Email: http://www.genome.clemson.edu/orders/index.html		

```

source
1. 430
/organism="Lycopersicon esculentum"
/cultivar="R11-13 (Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLES2F21"
/clone_lib="tomato susceptible, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLES - Tomato Pseudomonas Susceptible EST Library.
Directionally cloned cDNAs inserted into pBluescript SK(-)
) at 5' end with EcoRI and 3' end with XhoI site"
BASE COUNT      131 a      87 c      109 g      111 t
ORIGIN

```

Query Match	51.1%;	Score 24;	DB 10;	Length 438;
Best Local Similarity	61.4%;	Pred. No. 45;		
Matches 27;	Conservative	6;	Mismatches 11;	Indels 0;
			Gaps	0.
QY	1	tgtgtrgggcygacatawgtlyggttcaytaagccgyaaggygc	44	
	:		:	
	:		:	:
Db	345	TGTGTTAGGCGTTACATTTCCCATTTTCACTAAAGCAGCCAGGTCC	388	